

COMMENTS

SECURITY VERSUS DECEPTION IN PARAPSYCHOLOGY

By J. B. RHINE

From the simplest beginning in any science, precautions have to be taken to insure the product of the field against errors of observation, recording, logic, evaluation, reporting, and other uncertainties. In parapsychology this general concern over the basic security or reliability of the results divides conveniently into four major questions.

First: Have the experiments been firmly controlled against counterhypotheses (mainly *sensorimotor leakage*)?

Second: Have the *statistics* been appropriate?

Third: Would the problem *logically* permit a definitive conclusion if significant results were obtained?

Fourth (the topic of this paper): Has the research been adequately secure against *experimenter deception*?

The first and second of these major areas of insecurity were the leading counterissues to ESP in the 1930's. For example, the early test results at Duke were first suspected by their critics mainly of being produced either by sensory leakage or by improper handling of the statistical evaluation of the data. These were the main topics in the critical attacks made at the APA roundtable at Columbus in 1938. There, however, (along with the informal meeting of the American Institute for Mathematics at Indianapolis the year before) the concern about statistics was largely quelled, and by 1940, with *ESP After Sixty Years* (Rhine, Pratt, Smith, Stuart, & Greenwood, 1940) in print, the issue over sensory cues was fairly well resolved; the successful tests for precognition that followed definitely ruled out that problem.

By 1955, another of these areas of insecurity received major attention when Dr. George Price's critique of ESP appeared in *Science* (Price, 1955). His critical attack by-passed the first two issues and moved on to number four on the list, the reliability of the research personnel. Curiously enough, however, the reason Price gave for

going all-out on the issue of experimenter fraud was the number three issue on the above list—that the problem of the occurrence of ESP was not a scientific one. It simply could not qualify, as he saw it. (Nevertheless, with perfect inconsistency, he did propose a test design himself, an “adequate” one.)

But in the meantime, the questionnaire studies by Dr. Lucien Warner (1952, 1955) and by Warner and C. C. Clark (1938) were bringing out strong indications that American psychologists were not rejecting the ESP hypothesis as unscientific. In large majorities they were accepting the problem (that is, as distinct from the *answer*) as a legitimate one for psychology, but they were still hung up on the first two major questions, cues and chance.

Finally question number three has now been given its day. Parapsychologists themselves are at last beginning to re-examine critically in advance of actual research the scientific testability of their hypotheses in order to avoid the wasteful frustrations of the past. (I discuss this number three type of problem, with telepathy as an example, in an article to appear in this journal in June.)

THE GENERAL DECEPTION PROBLEM

In turning now exclusively to the question of deception, I am aware that it may seem late in the day to open up a general discussion of this problem. But the late timing is not so surprising. Students of parapsychology will recall Professor Henry Sidgwick's classic prediction that questions of the honesty of investigators would arise when all other counterexplanations had failed; he expected it to be a late issue. However, it is not because of any such desperate last-ditch status of the case for the occurrence of psi phenomena that the deception problem is being opened up here now.

Rather, the stimulus for making this review came from realizing how comparatively slow the recognition of parapsychology has been, and from reflecting over the possible factors responsible for this. It occurred to me that doubts may still exist in many minds which are too unclear and unsupported to be actually expressed, but sufficient to deter a more positive interest and reaction. As I thought of the more explicit attention that has been given to the other three major questions of research reliability listed above, it appeared quite possible that this more subtle, slightly distasteful, and sometimes em-

barrassing issue of fraud might need more frank and forthright recognition and response. It seemed, therefore, worth the effort to review the grounds for concern over this fourth major problem area, to outline what has already been done to cope with it, and then to show where the psi research field stands today with regard to it.

Subject Deception

To avoid confusion I will pass over the problem of deception by the test subject as belonging to question number one. In the early days when mediums and stage performers composed a large percentage of the participants in tests and demonstrations, the question of trickery applied mainly to the performers rather than to experimenters. Even as psychical research became more experimental, beginning in the 1870's, and the testing of subjects came under better control, one of the main purposes of the experimenter, of course, was to exclude all the possible deceptive (and other) practices of the subjects that would permit sensory cues. But as indicated, this exclusion of sensorimotor leakage comes under question number one. The more elementary problem of subject deception had to come first, and it has long since ceased to be a major issue.

Self-deception

Another order of deception, this one involving the experimenter himself, can also be passed over briefly here although it, too, has been an important one. This is self-deception, an expression that actually covers a range of relatively innocent mistakes. However, it parallels many of the types of deliberate trickery and is an important element in every science. I have personally seen more of this "fooling one's self" type than I have of conscious fraud. A few examples will suffice here.

Innocent, so-called "self-deception" is, of course, most likely in inexperienced observers, but it is possible even in highly trained individuals, especially those approaching parapsychology from some other field in which problems and controls are different. Even some experienced psychical researchers have been deceived for a time because of an excessive and disarming trust in, and attachment to, a test or overconfidence in a research assistant. Again, lack of experience with the dangers of defective test cards (or other test materials) or loose

test conditions may mislead the honest researcher, especially one who is working in isolation. But he can also innocently misuse his records, take liberties with his statistics, and most easily of all, make unwarranted interpretations of his results.

But let us pass over this nonfraudulent section of the experimenter-deception aspect of psi research because, first, it will be partly covered incidentally in the discussion of deliberate deception. Second, unintended errors of this kind are not likely to get by the editors of today without being detected since they result mainly in weaknesses covered by the first and second questions. They are in any case likely to occur in the work of comparative beginners; and even with the slightly improved chances of training that are possible today, the subspecies of unprepared experimenters may soon be approaching extinction.

WHAT ABOUT DELIBERATE DECEPTION BY EXPERIMENTERS?

Those psi researchers who have been at least suspected of being really crooked are all that are left for discussion; and in the final analysis, it can be said that this small but untrustworthy group is today all but threatened with extermination also. The known case histories go back somewhat to certain earlier stages of the research (or out to those that still need instruction in the elements of safeguarding). Let us see in brief review how reasonable this optimism appears.

Background

It is not necessary to look all the way back to the founding stages for untrustworthy psi researchers, since the weaker characters involved are not the kind to "rough it" over the first strenuous stretches of a rugged research road. They do not often appear on the scene until there is something else than the actual research results to be gained (something like easy notoriety). Also, I will stay within the more active *experimental* psi testing era that began in the 1930's, since, as I have said, the trickery in the heyday of mediumistic demonstrations almost entirely involved the subjects instead of the experimenters.

Even after my monograph, *Extra-sensory Perception* (1934), it was some years before the ugly hand of fraud began to appear

among the experimenters who followed up on the ESP work. There were, incidentally, numerous and various innocent errors that were more typical of the period. For example, one well-known mathematician came to my laboratory to try to convince me that my results were meaningless (as he put it, they were close to mean chance expectation). This, he said, was about 7.5 hits for a run of 25 trials with five target symbols; but when I saw his analysis it was possible to show him his misinterpretation of the conditions of the method. (He wrongly assumed feedback on each trial.) Another visitor, one of the leading psychologists of the day (1935), reported a negative (i.e., below-chance average) deviation of approximately 2.5 hits per run on 500 runs in which he tested himself. But he refused to submit a paper for publication and I could only infer that he made the same mistake with reference to the expected chance average that the mathematician had made, but that he was trying to *avoid* making hits. These men were perhaps a bit overeager to disprove the new claim, but I see nothing dishonest in these diligent efforts to "expose the error."

Perhaps there is a shade of difference in parallel cases of some individuals and departments of the great universities. One of these well-known departments of psychology invited me to give a report on the ESP work. A staff member there had conducted some successful ESP tests (unknown to me), but he did not mention them at all during my visit. When asked later by a fellow staff member why he had remained silent about them he replied, "Do you think I wanted Rhine to be able to go around saying our university had confirmed him?"

As a matter of fact, during the late 1930's there were many academic people, both staff and students, who independently confirmed the ESP tests of clairvoyance I had reported in 1934. (There were also many who, for one reason or another, failed in their attempts to repeat the tests, a fact that means little because there are so many wrong ways to do it.) Because of the hesitation on the part of many of the successful experimenters to publish their results, the mass of confirmation we might have been able to claim was never normally reported. A small part did come out anonymously; some was given restricted circulation by mail; and some of the best records (one described only orally by a well-known professor of psy-

chology) were deliberately destroyed to insure non-publication. Is not this a kind of deception-by-omission?

A Sampling of the Worst Stage

A certain change occurred, however, after World War II. Parapsychology had to some degree prevailed over its critics and had become almost excessively popular, sufficiently so to attract a number of the band-wagon type of "pioneers." In fact some of these enthusiasts claimed they had done research in the psi field and presented papers, either for publication or for a convention program. From these adventurers I have selected a dozen cases to illustrate fairly typically the problem of experimenter unreliability prevalent in the 1940's and 1950's. These twelve individuals themselves are all rather hard to classify. As it was, however, four of them were caught "red-handed" in having falsified their results; four others did not contest (i.e., tacitly admitted) the implications that something was wrong with their reports that seemed hard to explain and they did not try. In the case of the remaining four the evidence was more circumstantial, but it seemed to our staff they were in much the same doubtful category as the other eight.

What sort of people were these more obvious tricksters on the border or at least near the border of the field for a time? They ranged widely in many ways. Seven did not have the doctorate, although all were eager to get graduate degrees. Three of the seven were found to have claimed a degree fraudulently. Several were persons of evident ability but were located (some of them abroad) where research in parapsychology was extremely hard to manage but not nearly so hard to fake.

With all the worldwide publicity ESP research received so freely at that stage, it doubtless seemed easy to many weaker minds to concoct an experimental report based partly on their own imagination. In all the cases I knew personally, however, there was indication of some actual testing having been done. But this reaction was, with most of those who tried it (and were caught), shockingly crude and shameful. Perhaps a quarter to a third of them were able, clever people who need not have used trickery at all; they could surely have learned to do careful, effective testing. Odd as it may appear to some people, the ablest among this little collection of weak characters

seemed to be the most irresponsible. This observation, plus the fact that we could obviously not run a character-rehabilitation clinic, led us to discourage further contact at once.

Fortunately the culprits have thus far been caught (at least in our "known" cases) before serious damage has been done. Then, too, as time has passed our progress has aided us in avoiding the admission of such risky personnel even for a short term. As a result, the last twenty years have seen little of this cruder type of chicanery. Best of all, we have reached a stage at which we can actually look for and to a degree choose the people we want in the field. Finally, as will be seen in a few more pages, we have been able to do quite a lot to insure that it is impossible for dishonesty to be implemented inside the well-organized psi laboratory today. So after one further step into the background of the deception problem, I will be ready for the search for solutions.

Fine Points in Developing Methods

What makes this next step difficult is the fact that the indication of trickery described in this section is not nearly so definite in all cases as in the preceding one. (As a matter of fact, I shall even cite one case in which someone else claims evidence of fraud—a judgment with which I disagree. I shall use it partly because it is already in print, but mainly because it will illustrate my point no matter who is right about the charge of fraud.) We must deal firmly with all *possible* deception in parapsychology to make this problem the negligible consideration it has become in most other sciences.

One other qualification is needed. Some of the breaches of faith (suggested or proved as the case may be) in this further group of examples are so minor that I can be justified in making them out to be probable cases of culpable malpractice (as I aim to do only in principle) simply because this is science and because this particular science is in an inordinately sensitive stage.

The members of this second group were all better qualified for psi research than those of the preceding selection (the "dozen"). They all knew the rules and standards that had been developed through the years, standards which compared favorably with those of the neighboring sciences. In fact, it was these more advanced test procedures that had largely ruled out the earlier types of fraud dis-

cussed above. Thus the kind of deception left as at least a possibility at this more sophisticated stage mainly consisted of ways of somehow lowering or somewhere skirting these precautionary bars in some slight degree and thereby leaving the safeguarding doors ajar by as much as a tiny crack or more. This borders in some cases on little more than a reasonable suspicion by someone of intent to cheat, leaving in most instances a possible alternative explanation. However, psi is so important and so revolutionary that it seems reasonable to aim at allowing no possible opportunity for an experimenter to mislead even a little.

Example No. 1. One of the most frustrating weaknesses among research workers has been the difficulty one experimenter has in actually exercising the precautions against dishonesty when collaborating with his trusted friends (and most of all, with close relatives). This difficulty is more acute among people inexperienced with regard to experimental deception, especially when dealing with colleagues whose relationship has been friendly and of long standing. In example No. 1, two mature experimenters had undertaken to do a well-designed double-blind experiment in which psi test data were to be correlated with another series of non-psi measurements. The main control over experimenter insecurity lay in avoidance of any inter-experimenter leakage of the individual scores until the final well-guarded checkup. However, the two experimenters in this case were discovered to be covertly exchanging tips with each other whenever outstanding subjects turned up in the series. These experimenters doubtless believed they were sufficiently objective not to need stringent rules. They were, of course, too interested in the results to wait until the series was finished.

Was this really cheating? Perhaps a sufficient answer can be found in the fact that it was done surreptitiously. I do not need to say (or wish to imply) that they actually *were* biased by this leaked information in their final analyses, but at least the results were not approved for publication and the individuals were not encouraged to continue work at the center.

Example No. 2. This case is essentially similar except that only one of the two experimenters was suspect. The experimental design again was of the double-blind type, and again analyses were to be made of the correlation of ESP test data with another set of mea-

surements of a non-psi type. In principle the double-blind design could be perfected to a high degree because of the distinctiveness of the two types of measures used. These two sets of data were both meant to be analyzable on wholly objective (double-blind) lines. One experimenter (E-1) was responsible for the psi records and the other (E-2), for the non-psi recordings. The results of the correlations were quite significant throughout a number of confirmatory repetitions and with several E-2's as well as with a variety of subjects. E-1, however, was the single common factor throughout.

The point of interest is that when the exchange of recordings was made there was a short gap in the double-blind coverage in which E-1 briefly had sole possession of both sets of records (before delivery to the analyst doing the blind checking). Also, there were some adjustments of timing between the two series of data that could with inspection somewhat influence the evaluation. Here was a gap that needed to be closed to prevent any possible manipulation of the timing.

When it was arranged that E-2 would leave no time lag at all for E-1 to have unwitnessed possession of both records, the successful performance discontinued. It resumed, however, when the original conditions were restored; then it stopped when the time gap was closed again by E-2—with all else kept the same.

What was wrong here? Everyone urged going on to see what emerged with further patient variation, everyone but E-1 himself; he left, and fortunately the experimental results had not been published so that no one was misled by this particular instance.

Example No. 3. One of the better controlled psi test methods consists of the guessing of card (or symbol) order as a way of predicting future events, i.e., precognition. In checking these guesses against the future random targets, two experimenters can safely use double-blind conditions to insure against error. Much of the best psi work has been done with this method, with two responsible experimenters and with independent records to be matched jointly (or better still, matched independently with the use of duplicate records). But unfortunately it has not always been done in the way the design requires. Example No. 3 is a case in point.

In this experiment E-1 supervised the test performance of the subject and arranged for E-2 to be ready to prepare the future target

series of symbols independently as soon as the subject's records were completed. The method thus in principle provided strong protection against any possible cheating by the subject. When properly conducted, it also guaranteed two sets of independent records that neither experimenter could interfere with by trickery.

In the editing of a report of this experiment special analyses were made of the data that showed an interesting hit distribution on the record sheets; this in turn suggested a further investigation of the actual test conditions, and this revealed a rather simple trick. A few spaces on E-1's hand copy of the subject's calls were left blank (as though by accident) until the actual checkup when they could be filled in as hits by E-1 himself. The use of duplicate sets of records to be exchanged by E-1 and E-2 at the checkup time had been omitted, evidently by E-1's intention. Completely mutual vigilance in the joint checking procedure was also obviated. With a well-trained and more watchful E-2 on the job, this cheating could not have occurred.

Example No. 4. Some of the best test methods in the ESP researches have been the gamelike card-matching techniques. They went through several forms, one of them eventually taking shape as STM or Screened Touch Matching. It was used most often in clairvoyance tests and finally evolved into a procedure in which the five key cards were hung on the subject's side of an opaque screen with a one-way opening which afforded visibility on E-2's side. By using a pointer visible to E-2 through the opening, the subject could indicate his guesses for the target cards being laid down one-by-one. Thus the subject could not see the test cards in E-2's hands, and E-2 could not tell what the key-card order was; that is, the keys were to be rearranged as randomly as possible by the subject before each run, under the continual observation and collaboration of E-1. On the opposite side of the screen, E-2 shuffled the target deck for each run.

The best known work with STM was the Pratt and Woodruff experiments (1939). Since the question of experimenter deception regarding this work has already been raised in publication by C. E. M. Hansel (1961), I can use it as case No. 4, and do so without accepting the theory of fraud; this interpretation was definitely not proved by Hansel. However, the mere fact that the actual conduct of the

experiment was such that trickery was a conceivable possibility qualifies it for discussion here. There is some virtue in considering the possible weakness of the method whether or not any advantage was taken of it (a policy I also follow in case No. 1).

To support his charge of fraud in the Pratt-Woodruff tests, Hansel presented the results of analyses of the records of the highest-scoring subject in the experiment and claimed that E-2 could sometimes have partially identified one or more of the five key cards (partly because of inadequate rearrangement of them by the subject) and also that the records of the hit distribution of this one subject supported the hypothesis that E-2 could intentionally have taken advantage of this knowledge in laying down the test cards opposite the key cards (to some extent known to him). This was not proof that E-2 necessarily did this, and Pratt and Woodruff (1961) pointed out serious errors in Hansel's argument for his fraud hypothesis. A further round of the analyses begun by Hansel has been extended by George Medhurst and Christopher Scott to a larger group of subjects in this experiment and is awaiting publication in the *Journal of Parapsychology*. The argument will then be continued with a further rejoinder by Pratt. This research can serve meanwhile to illustrate a step in the long and tenuous effort at improving experimenter security through developing experimental design.

The point that is specifically relevant to the present issue is that what was in 1939 considered a strong test design, i.e., the two-experimenter, double-blind technique, was not as secure as it could be made. It was a clear advance over general psychological test conditions and had not at that stage met with criticism either in or outside of the Laboratory. Nevertheless, modifications followed, even in the next year's research, that tightened the precautions further; for example, in precognitive matching tests (Rhine, 1941 a).

Other Points Needing Attention

I will end this section with a few added generalizations about these finer points of experimenter security. Cases could be drawn, for example, from projects in which a too free abuse of statistics has lent itself to deceptive conclusions. While the help of statistical editors on the *Journal of Parapsychology* has been a most important service against this hazard, one gap that has been hard to keep closed

is the omission (on improper grounds) of data that might legitimately belong in a report. Such cases are often impossible to rate because of the lack of full records or because of a question as to the completeness of the reporting; but whenever the decision is found to have been a secret one (that is, without an objective sharing with colleagues qualified to judge) that makes a serious difference. If there were need (and space for more illustration) there would be an example of this type. The subtle, private judgments about what data to "declare" in reporting constitute an area that needs the fullest possible safeguarding.

A final risk to be listed will surprise most of those accustomed to the justly exuberant confidence inspired by the reassuring words "automated," "electronically recorded," "computer analyzed," and the like. While great benefits are already being contributed by the advanced technology now available (or at least borrowable) in some laboratories, and the hopes of almost endless further gains are high, some caution definitely needs to be extended here too, even though on a slender basis of judgment as yet. An early experience of my own with fraud (in the case of Margery, the Boston medium) showed me rather convincingly that apparatus can sometimes also be used as a screen to conceal the trickery it was intended to prevent. Some of the suspected instances of intentional selection of data already mentioned were not necessarily insured against just because they had been more easily or automatically recorded. Perhaps the main thing to keep in mind is that the increased reliance on the equipment (which makes double-blind methods based on two experimenters appear less necessary, almost a luxury at first thought) does not necessarily mean that the need of the two-experimenter design is in any way supplanted; the experimenter-machine team is not at all the same as the two-experimenter plan. It would be a real retreat to think it ever could be. Machines will not lie, but . . .

The general point of this section is that in a developing research field, the methods themselves are always on trial from the first. *If we have to argue* over the adequacy of the design or the trustworthiness of the experimenter (or the subject) it is wise to back up and improve the method before advancing further or expecting really firm credence from fellow scientists. The emphasis has to be on the tightening of security.

SAFETY MEASURES AGAINST DECEPTION

What can now be said (in outline) of the way to judge whether or not a psi experimenter has been reliable? The primary effort, of course, has been directed into seeing that dishonesty would be impossible if significant results were obtained by the experimenter under the prescribed conditions.

One of the earlier steps taken was the use of two sets of independent records (the record of the targets by the experimenter and the record of the calls by the subject), both to be handed to a second experimenter before checking was begun. This method was used in our laboratory at Duke in 1933 in the Pearce-Pratt series of clairvoyance card-guessing tests (Rhine, 1934). In addition, the two-experimenter practice began shortly thereafter in a later subseries of the same experiment. This meant double witnessing at the experimenter's end of the test, with the subject being sent to another building. Another precautionary step reported by Pratt and M. M. Price (1938) was to have two experimenters testing individual subjects for clairvoyance, with one experimenter handling the screened cards and the other recording the subject's guesses; the two experimenters then checked the results together. By 1939 the two-experimenter STM method was developed, as described above in the Pratt and Woodruff (1939) experiment. A further advance of the double-blind card matching method was reported in the 1941 *Journal of Parapsychology*. For example, completely independent (double-blind) checking of records was introduced in both of my reports of that year, one on clairvoyance with the target cards in sealed boxes (Rhine, 1941 b) and the other in precognition matching tests (Rhine, 1941 a). In addition, the latter report was done with the two experimenters operating completely double-blind, one handling the target records in one room; the other, those of the subjects in a second room. As conducted and reported, these conditions were, I think, very secure; and I know of no criticism of them thus far.

Still other variations of two-experimenter and double-blind techniques were introduced in the years that followed, especially with the precognition tests. The latter became the best controlled of all the types of psi known and tested up to that time. For example, with two trained experimenters adhering to the rules, neither experi-

menter could deceive the other; that is to say, it looked as if deception would have to involve *collusion between the two*—a rare situation indeed. (My No. 1 example of two-experimenter deception, or at least dishonest breaking of the rules, is the only case in my memory.) But, as I have already indicated, a good method itself is no guarantee that it will always be faithfully carried out. One over-tolerant experimenter alone can innocently allow another to get around the barriers that have been set up to insure the reliability of both. What must be kept in mind is the possibility of the method itself being less than cautiously applied and the necessity of preventing that from occurring unnoticed. These few examples will perhaps illustrate the need to continue to reinforce experimental design in parapsychology still further against possible experimenter unreliability even after all the controls developed thus far have been taken fully into account.

INCIDENTAL EVIDENCE: THE "SIGNS OF PSI"

Yet, even with all I have said, let us remember, for balance on this difficulty, that most other branches of science have already matured to the point where the problem of experimenter trickery causes no great concern. That is partly because deliberate fraud would be too quickly spotted and exposed at their present stage. Also, in the more advanced sciences the research personnel is increasingly well selected through a long program of university training. But it was not always thus; I recall that fifty or more years ago, there were notorious cases of experimenter fraud in physics, biology, and medicine, among other fields. Obviously the possibility of easy repetition of tests as a way to check up on a new claim offers the best protection against trickery in research. Parapsychology has only in very recent years been coming into the stage of a reasonable likelihood of confirmation in other laboratories. Such repetition naturally requires first of all that these other laboratories exist, and again that they have staff members qualified and equipped to repeat the new findings. This period in parapsychology is only beginning, and is coming along slowly at that.

However, we have at least got past the older phase of having to use detectives and magicians to discover or prevent trickery by the subjects. The psi laboratory's experimental precautions of early

years were (and had to be) mainly countertrickery measures (directed especially against subject fraud), although they had to safeguard against innocent mistakes as well. All of these devices were defensive in character and were quite necessary until enough knowledge was acquired to bring parapsychic phenomena into good laboratory test conditions. Then it became possible to look beyond the mere evidence for more *positive* "signs of psi," incidental earmarks of a more distinctive and peculiar nature. A few representative selections of these signs will be reviewed as an important part of the answer to the deception question reached thus far.

Decline Curves

One of the first of these "signs" came from the work of my first graduate student in parapsychology, H. L. Frick (Rhine, 1934). In his exploratory experiments in clairvoyance for the A.M. thesis he made a series of daily runs of 100 guesses of playing-card suits. The pooled results were close to the expected chance average, but later examination showed a continuous decline of the average scores over the five, 20-trial segments of the run. The fifth (or last) segment averaged about as far below the expected chance mean as the first one did above, and the two were significantly different. When this type of decline was found to recur frequently in later analyses of comparable test data in other researches, it began to acquire some useful identification value. In fact, it was sometimes the only safe evidence of psi to be found in a given experiment.

This decline in the run thus became a "sign of psi"; it could serve as evidence against experimenter deception when it was discovered later by another analyst. Such a finding can be about as objective a type of evidence as fingerprints. In forming my own judgment about the reliability of psi research, I have leaned most heavily on such hidden evidence, mainly based on significant internal differences. The analysis usually has the great merit of being completely repeatable. To appreciate it fully, however, one needs to understand the evidence and nature of the various other position effects, as well as the related psi-missing effect. I have already reviewed elsewhere the main evidence that has accumulated on these effects (Rhine, 1969 a, 1969 b).

The decline curves reached their greatest value for security to date

in the quarter-distribution (QD) analyses of the PK research data resulting from tests with dice. In these studies more than 30 years ago, Dr. Betty M. Humphrey and I made QD analyses of hit distributions over the record page (Rhine and Humphrey, 1944 a) in all the available PK test records for the preceding nine-year period. None of the 18 experiments available had been conducted with a decline of scoring rate in mind; yet, taken as a block, this great mass of research data showed a highly significant diagonal decline in the right-and-downward direction. This diagonal decline was conclusively confirmed by a further internal consistency test made on smaller record units (sets) within the page (Rhine and Humphrey, 1944 b). Moreover, an independent recheck was conducted by Dr. J. G. Pratt (1944), and this, too, produced almost perfect confirmation. To cap it all, a published invitation for still another analysis was made; it has thus far, after a period of three decades, received no takers. It appears to me to be the firmest block of evidence yet offered in a behavioral science in support of a new hypothetical principle.

The U-curve

If this were a book, I would go on next to develop chapters on the various other signs of psi to be derived from position effects as definitive types of evidence against experimenter deception (and error in general). One of these would be devoted to the U-curve effects. The U-curve was produced as another incidental sign of psi which came out of the use of one of the early card-guessing tests, the one called DT (down through) (Rhine, 1934). In these tests the subject made 25 responses (usually written) to guess the card order *down through* the undisturbed deck. By the time of my study of comparative test techniques (Rhine, 1941 b) it was well known that it was the *subject's response to the structure of the record sheet* that produced the U-curve. The sheet which he used had 25 spaces, in segments of 5 each. The hits tended to fall into U-shaped curves of frequency in a lawful way that could be statistically evaluated. A "salience ratio" statistic was devised by Dr. J. A. Greenwood (1941). The U-curve, like the decline curve, was a good antifraud feature when it emerged in the second stage of analysis, especially from test data of an experiment in which such curves had not been anticipated by the researcher.

Psi-missing

The psi-missing effect has furnished some of the surest signs of psi, evidence that meets the highest standards of objective science. To give this topic a better starting point I would go back to my precognition article (Rhine, 1941 a). The subjects in these card-matching tests were kept in separate groups of children and adults. The adult results gave deviations that were *below* chance to almost the same extent that the deviations of the children were *above* chance; both gave larger deviations in reward sessions, and smaller ones (about half as big) in sessions when no rewards were given. The experimenters had no basis then for anticipating such lawful distinctions as these (and other) breakdowns in this two-experimenter, two-room, double-blind experiment. The point is that natural lawful psychological effects occurring to the adults in these negative deviations paralleled the opposite (positive) results in children, and this fact goes far toward carrying the rational mind beyond the point of doubt about the trustworthiness of the experimenters.

It would require many chapters to round up the peculiar but lawful effects of psi-missing, some of them even more intricate than the differential signs of deviation shown by adults and children in the research just cited. These complex findings have often been completely unanticipated by the experimenters and thus have allowed no possible opportunity for deception by them.

For my own part, I have most effectively reduced my own skepticism by watching decades of these signs of psi emerge, often with such surprise as to make the experimenter himself an obviously "innocent bystander." Those who are not familiar with this most solid of psi evidence can follow it up through the past records or in its ongoing development. No critic, so far as I know, has tried to do this yet. But the historians are coming!

And yet, hard as it is to say it, I am not entirely satisfied today with all these many objectively factual and repeated signs of psi with their various types of internal verification, appealing as they are to my own rational scientific judgment. They do most forcefully reassure me personally; and yet, for many more than myself I think something more is needed for the field. For one thing, such extensive, involved analyses as those of the QD studies cannot be made by

just anyone and could not be made so well again without that decade of accumulated records of PK data which were on hand in 1944.

Then, too, these signs of psi are puzzling in themselves. We can easily use them as effective empirical test devices and very often have done so. More than that, they pose challenging ideas about how psi functions. But for security purposes the signs have had their best safeguarding value when completely unexpected as effects by the experimenter, and thus not possibly attributable to his own intentions, conscious or unconscious. Therefore, as evidence against fraud, they lose some of their potency as they acquire familiarity.

Rather, there ought to be ways of so well embedding adequate safeguards against experimenter deception in psi research methodology that no reasonable question of the honesty of the researcher will ever arise. I think these safeguards can be provided along the lines suggested in the section to follow.

A PROGRAM FOR EXPERIMENTER SECURITY

First, let us recall why there is concern about fraud in parapsychology, as compared, for example, to fraud in chemistry. As already stated, it is because in psi research it is still harder to test a new claim than it is in most sciences, and there are fewer researchers in a position to do it. This suggests, then, an obvious need, not only to encourage more widespread effort at repetition, but also (as a new emphasis) to encourage every psi researcher to make his experiment as easy to repeat as possible. If from the start he recognizes independent confirmation to be an essential part of his own goal, he will be able to do much to aid and insure such replication. Exchange of information and even visits with other research workers, loans of equipment, subjects, and the like are all advantageous in extending research into other laboratories for duplication by other experimenters. Logically enough, for parapsychology's present stage these *independent confirmations take on a value rating above that* of the original work itself simply because of the assurance they give on the security problem. Because of this current situation in parapsychology any new piece of work should be taken as almost a sort of pilot research. In order to give it the proper status of acceptability it is more than ordinarily important for another laboratory to repeat it with adequate success. This then will complete the project as a

sufficiently effective research contribution. Is that too high a standard to set for our field? Not for a science that is still so obviously fighting its way to the acceptance required to render it useful and meaningful.

The greatest difficulty will be in obtaining the cooperation necessary for the large-scale repetition needed. As it is, there is far too little "will to repeat" in this field at present; most researchers want to be innovators, since it looks more creative. But those who want the field to be taken seriously beyond its own small group will in time see the basic need of this reinforcement of security through repeating each other's experiments. Obviously we are all in it together.

Second, a similar step toward reinforcing the reliability of the individual experimenter's role in the research would be to include one or more coworkers as early in a research series as is feasible. This would also give further assurance of success when the transfer of the project to another center is undertaken; i.e., to show that the experiment is not dependent on one experimenter alone is best managed right "at home," and as early as possible. In principle, the more teamwork the better, with the idea in mind of extending the number of experimenters who can share in the responsibility for the reliability of the conclusions reached.

At the same time, this idea of teamwork in research needs careful study, planning, and, of course, the necessary facilities. It is a painful fact that few places are ready yet for *all* these needed advantages. Also, individual initiative needs to have its place in a research field that is to retain its fertility. No one's personal enthusiasm must be dampened by an overspreading of responsibility. However, I have seen in several fields and in more than one center enough examples of successful research-group life to make it reasonable to hope that parapsychology can achieve the full benefits of teamwork, security among them. The idea largely reduces to this: a good psi research worker can multiply (not merely increase) the value of his contribution by use of a well-developed, suitably staged partnership. Such a partnership can be one that shares concern and responsibility for basic precautions and makes the eventual transfer of independent replication to other experimenters easier and much more likely to succeed.

Third, another strong fence against personal unreliability can be

built by developing the best possible system for the exchange, the registration, and the safe preservation of research data. It is recognized that the research worker must be assured reasonable independence in order to cultivate and shelter his own ideas at the sensitive stage of innovation. He should be free to do his own preliminary explorations within the field more or less as he prefers; but when he has performed a promising pilot experiment and wants to set up a confirmatory project, he then needs to go on record with his group and to try openly to share his project with one or more of his colleagues. The research should go through one round of experimental confirmation after another with the *center's review system keeping the complete records*. This is an easier matter today with modern equipment; but it can always be done in any set-up, and it is necessary in order to avoid risk of omissions and improper selection of reportable data. With frequent reviews at staff meetings, and (as the work grows) at suitable conventions, the step-by-step developments will be shared and welcomed with growing interest beyond the original laboratory. All this sharing of progress (as well as failures) can do something not only to sustain morale but also to keep the data record straight, complete, and always ready for review and re-examination.

Fourth, in order to be taken seriously, the research of an experienced worker should be aimed well beyond a single initial experiment. When the immediate experiment is a recognized start on a long-view objective, it carries added assurance, security, and magnitude of purpose *because of* this larger perspective. The greater the problem to which the experiment makes a relevant approach, the more conviction the results are likely to carry. Moreover, the more closely the immediate project relates to already established territory, either within parapsychology itself or other branches of science, the more substantial and well-based it appears and the more trust it inspires in the credibility of the experimenter. The growing interrelations in the emergent picture of the nature of psi rank high in the building of a requisite overall confidence. This, of course, is the long-view answer to all the doubts about the field of parapsychology.

Finally, the really best clincher of all can well come with the finding of linkages between a new result and one or more of the identifying signs of psi, especially when the linkage is such that it could

not possibly have been anticipated and "planted" by the original experimenter. These linkages will always be prime evidence, even if only incidental to the more adaptable program I have outlined.

As we proceed now to close in on this experimenter-deception problem by combining these added safeguards with those already in use, I do think it should be possible almost immediately, if a resolute move is made, to put an end to the long-lingering anxiety that I think has sapped the confidence our field has needed for its effective recognition.

As distrust of experimenter security diminishes and the concern over personnel unreliability reduces to the level of the older sciences, the findings of the psi research field should, after having waited for too many generations, come to be accepted on their actual objective merits. That is enough to ask, but nothing less than that should any longer be considered acceptable.

A FINAL REFLECTION

Parapsychology appears to those who know its history to have come a rather long way in its advances in methods, even if it has taken a long, long time to do it. It has come far in ferreting out all the many far-ranging doubts and questions and counterhypotheses both its friends and foes could identify or even merely suspect. It has stretched out its patient pursuit of an ever more conclusively tight experimental design and statistical evaluation until the growing burden on the research field is almost frustratingly depressing to much of its personnel. If now there is still a latent last-ditch sort of hold-out against it, this may be due in large part to a residue of suspicious, half-concealed distrust of the human researcher that keeps parapsychology in a state of futile unacceptability—a state in which scientists distrust their own best methods.

If this uneasiness over unreliable personnel (no matter how rare it is) is indeed what has largely been sustaining the existing hesitancy over parapsychology so long after it has had twice the normal period of trial, by all means let us firmly insist on a fuller sharing, controlling, and accounting of the entire research operation. This added vigilance need not impose any undue burden on research when it becomes adopted practice. Rather, it should relieve the active concern one often feels—but always hesitates to express—when impres-

sive results occur, especially in some other laboratory. In other words, now that we have found out how to verify the occurrence of psi, let us try in the way we do the research to make it completely convincing.

REFERENCES

- GREENWOOD, J. A. The statistics of salience ratios. *Journal of Parapsychology*, 1941, 5, 245-49.
- HANSEL, C. E. M. A critical analysis of the Pratt-Woodruff experiment. *Journal of Parapsychology*, 1961, 25, 99-113.
- PRATT, J. G. A reinvestigation of the quarter distribution of the (PK) page. *Journal of Parapsychology*, 1944, 8, 61-63.
- PRATT, J. G., & PRICE, M. M. The experimenter-subject relationship in tests for ESP. *Journal of Parapsychology*, 1938, 2, 84-94.
- PRATT, J. G., & WOODRUFF, J. L. Size of stimulus symbols in extra-sensory perception. *Journal of Parapsychology*, 1939, 3, 121-59.
- PRATT, J. G., & WOODRUFF, J. L. Refutation of Hansel's allegation concerning the Pratt-Woodruff series. *Journal of Parapsychology*, 1961, 25, 114-29.
- PRICE, G. R. Science and the supernatural. *Science*, August 26, 1955.
- RHINE, J. B. *Extra-sensory Perception*. Boston Society for Psychic Research, 1934. (Republished: Bruce Humphries, Branden Press, 1973.)
- RHINE, J. B. Experiments bearing upon the precognition hypothesis. III. Mechanically selected cards. *Journal of Parapsychology*, 1941, 5, 1-57. (a)
- RHINE, J. B. Terminal salience in ESP performance. *Journal of Parapsychology*, 1941, 5, 183-244. (b)
- RHINE, J. B. Position effects in psi test results. *Journal of Parapsychology*, 1969, 33, 136-57. (a)
- RHINE, J. B. Psi-missing re-examined. *Journal of Parapsychology*, 1969, 33, 1-38. (b)
- RHINE, J. B., & HUMPHREY, B. M. The PK effect: special evidence from hit patterns. I. Quarter distributions of the page. *Journal of Parapsychology*, 1944, 8, 18-60. (a)
- RHINE, J. B., & HUMPHREY, B. M. The PK effect: special evidence from hit patterns. II. Quarter distributions of the set. *Journal of Parapsychology*, 1944, 8, 254-71. (b)
- RHINE, J. B.; PRATT, J. G.; SMITH, B. M.; STUART, C. E.; & GREENWOOD, J. A. *Extra-sensory Perception After Sixty Years*. New York: Henry Holt, 1940. (Republished, Boston: Bruce Humphries, 1966.)
- WARNER, L. A second survey of psychological opinion on ESP. *Journal of Parapsychology*, 1952, 16, 284-95.

WARNER, L. What the younger psychologists think about ESP. *Journal of Parapsychology*, 1955, 19, 228-35.

WARNER, L., & CLARK, C. C. A survey of psychological opinion on ESP. *Journal of Parapsychology*, 1938, 2, 296-307.

Institute for Parapsychology
College Station
Durham, N. C. 27708

AUTHOR'S NOTE—At the page-proof stage of this paper on deception I received a copy of a recent article which I think it important to mention here. (I acknowledge that my acquaintance with it is due to Dr. Hans Kreidler.) The article, entitled "Pitfalls in Research: Nine Investigator and Experimenter Effects," is by Dr. Theodore X. Barber of the Medfield Foundation, Medfield, Massachusetts, and is Chapter II in a book entitled *Second Handbook on Research on Teaching*, edited by R. M. W. Travers (Rand McNally, Chicago, 1973).

Dr. Barber's article is a broadly oriented treatment of possible research errors and is admirably inclusive and systematic. While it is illustrated in terms of general psychology, it is quite relevant to parapsychology as well. Two of the nine pitfalls listed are closely related to this paper on deception. (Barber uses the term "fudging.") And a third danger point applies to what I am calling, in my paper on telepathy in the next issue of the *Journal of Parapsychology*, the choosing of solvable problems. It is particularly timely with regard to our own field to see this strong initiative in general psychology toward safeguarding the quality and security of research.—J.B.R.